AD NO. ___________________________
DTC PROJECT NO. 8-CO-160-UXO-021
REPORT NO. ATC-9409

STANDARDIZED

UXO TECHNOLOGY DEMONSTRATION SITE

BLIND GRID SCORING RECORD NO. 806

SITE LOCATION:
U.S. ARMY YUMA PROVING GROUND

DEMONSTRATOR:
U.S. GEOLOGICAL SURVEY (USGS)
BOX 25046, FEDERAL CENTER, M.S. 964
DENVER, CO 80225-0046

TECHNOLOGY TYPE/PLATFORM:
TMGS MAGNETOMETER/TOWED ARRAY

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

MAY 2007

Prepared for:
U.S. ARMY ENVIRONMENTAL COMMAND
ABERDEEN PROVING GROUND, MD 21010-5401

U.S. ARMY DEVELOPMENTAL TEST COMMAND
ABERDEEN PROVING GROUND, MD 21005-5055

DISTRIBUTION UNLIMITED, MAY 2007.

DISPOSITION INSTRUCTIONS

CSTE-DTC-AT-SL-E

MEMORANDUM FOR RECORD

SUBJECT: Operations Security (OPSEC) Review of Paper/Presentation

1. The attached record entitled "The Standardized UXO Technology Demonstration Site Blind Grid Scoring Record No. 806" dated May 2007 is provided for review for public disclosure in accordance with AR 530-1 as supplemented. The scoring record is proposed for public release via the internet.

2. I, the undersigned, am aware of the intelligence interest in open source publications and in the subject matter of the information I have reviewed for intelligence purposes. I certify that I have sufficient technical expertise in the subject matter of this report and that, to the best of my knowledge, the net benefit of this public release outweighs the potential damage to the essential secrecy of all related ATC, DTC, ATEC, Army or other DOD programs of which I am aware.

Michael Karwatka		May 2007
NAME (Printed)	SIGNATURE	DATE

CONCURRENCE:	NAME (Printed)	SIGNATURE	DATE
Program Mgr/Customer (If not ATC owned technology)	Patrick McDonnell		21 June 07
Directorate Director	Charles Valz		6/12/07
Directorate OPSEC QC and Team Leader	William Burch		22MAY07
ATC OPSEC Officer/ Security Manager	Jenell Bigham		27 June 2007
Public Affairs Specialist	Karen Dern		28 June 2007
Technical Director, ATC	John R. Wallace		29Jun07
(Return to ATC PAO for further processing)			
DTC GO/SES	N/A		

Encl
as

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
May 2007	Final	8 through 12, 15 through 17 May 2006

4. TITLE AND SUBTITLE
STANDARDIZED UXO TECHNOLOGY DEMONSTRATION SITE BLIND GRID SCORING RECORD NO. 806 (U.S. GEOLOGICAL SURVEY, TMGS MAGNETOMETER/TOWED ARRAY)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Karwatka, Michael

5d. PROJECT NUMBER
8-CO-160-UXO-021

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Aberdeen Test Center
ATTN: CSTE-DTC-AT-SL-E
Aberdeen Proving Ground, MD 21005-5059

8. PERFORMING ORGANIZATION REPORT NUMBER
ATC-9409

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Environmental Command
ATTN: IMAE-ATT
Aberdeen Proving Ground, MD 21005-5401

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Same as Item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution unlimited.

13. SUPPLEMENTARY NOTES
None

14. ABSTRACT
This scoring record documents the efforts of the U.S. Geological Survey to detect and discriminate inert unexploded ordnance (UXO) utilizing the YPG Standardized UXO Technology Demonstration Site Blind Grid. This Scoring Record was coordinated by Michael Karwatka and the Standardized UXO Technology Demonstration Site Scoring Committee. Organizations on the committee include the U.S. Army Corps of Engineers, the Environmental Security Technology Certification Program, the Strategic Environmental Research and Development Program, the Institute for Defense Analysis, the U.S. Army Environmental Command, and the U.S. Army Aberdeen Test Center.

15. SUBJECT TERMS
U.S. Geological Survey, USGS, UXO Standardized Technology Demonstration Site, YPG, Blind Grid, TMGS Magnetometer/Towed Array, MEC

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(Include area code)*
Unclassified	Unclassified	Unclassified	SAR		

ACKNOWLEDGMENTS

Authors:

Rick Fling
Aberdeen Test Support Services (ATSS)
Sverdrup Technology, Inc.
U.S. Army Aberdeen Proving Ground (APG)

Christina McClung
Aberdeen Data Services Team (ADST)
Logistics Engineering and Information Technology Company (Log.Sec)
U.S. Army Aberdeen Proving Ground

Contributors:

Matthew Banta
Aberdeen Test Support Services
Sverdrup Technology, Inc.
U.S. Army Aberdeen Proving Ground

William Burch
Michael Karwatka
Military Environmental Technology Demonstration Center (METDC)
U.S. Army Aberdeen Test Center (ATC)
U.S. Army Aberdeen Proving Ground

Patrick McDonnell
Booz Allen Hamilton (BAH)
U.S. Army Environmental Command (USAEC)
U.S. Army Aberdeen Proving Ground

(Page ii Blank)

TABLE OF CONTENTS

PAGE

SECTION 4. TECHNICAL PERFORMANCE RESULTS

PAGE

SECTION 5. ON-SITE LABOR COSTS

SECTION 6. COMPARISON OF RESULTS TO BLIND GRID DEMONSTRATION

SECTION 7. APPENDIXES

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of unexploded ordnance (UXO) require testing so that their performance can be characterized. To that end, Standardized Test Sites have been developed at Aberdeen Proving Ground (APG), Maryland and U.S. Army Yuma Proving Ground (YPG), Arizona. These test sites provide a diversity of geology, climate, terrain, and weather as well as diversity in ordnance and clutter. Testing at these sites is independently administered and analyzed by the government for the purposes of characterizing technologies, tracking performance with system development, comparing performance of different systems, and comparing performance in different environments.

The Standardized UXO Technology Demonstration Site Program is a multi-agency program spearheaded by the U.S. Army Environmental Center (USAEC). The U.S. Army Aberdeen Test Center (ATC) and the U.S. Army Corps of Engineers Engineering Research and Development Center (ERDC) provide programmatic support. The program is being funded and supported by the Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP) and the Army Environmental Quality Technology Program (EQT).

1.2 SCORING OBJECTIVES

The objective in the Standardized UXO Technology Demonstration Site Program is to evaluate the detection and discrimination capabilities of a given technology under various field and soil conditions. Inert munitions and clutter items are positioned in various orientations and depths in the ground.

The evaluation objectives are as follows:

a. To determine detection and discrimination effectiveness under realistic scenarios that varies targets, geology, clutter, topography, and vegetation.

b. To determine cost, time, and manpower requirements to operate the technology.

c. To determine demonstrator's ability to analyze survey data in a timely manner and provide prioritized "Target Lists" with associated confidence levels.

d. To provide independent site management to enable the collection of high quality, ground-truth, geo-referenced data for post-demonstration analysis.

1.2.1 Scoring Methodology

a. The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver-operating

characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}), and those that do not correspond to any known item, termed background alarms.

b. The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the blind grid RESPONSE STAGE, the demonstrator provides the scoring committee with a target response from each and every grid square along with a noise level below which target responses are deemed insufficient to warrant further investigation. This list is generated with minimal processing and, since a value is provided for every grid square, will include signals both above and below the system noise level.

c. The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such and to reject clutter. For the blind grid DISCRIMINATION STAGE, the demonstrator provides the scoring committee with the output of the algorithms applied in the discrimination-stage processing for each grid square. The values in this list are prioritized based on the demonstrator's determination that a grid square is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For digital signal processing, priority ranking is based on algorithm output. For other discrimination approaches, priority ranking is based on human (subjective) judgment. The demonstrator also specifies the threshold in the prioritized ranking that provides optimum performance, (i.e. that is expected to retain all detected ordnance and rejects the maximum amount of clutter).

d. The demonstrator is also scored on EFFICIENCY and REJECTION RATIO, which measures the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. EFFICIENCY measures the fraction of detected ordnance retained after discrimination, while the REJECTION RATIO measures the fraction of false alarms rejected. Both measures are defined relative to performance at the demonstrator-supplied level below which all responses are considered noise, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

e. All scoring factors are generated utilizing the Standardized UXO Probability and Plot Program, version 3.1.1.

1.2.2 Scoring Factors

Factors to be measured and evaluated as part of this demonstration include:

a. Response Stage ROC curves:

(1) Probability of Detection (P_d^{res}).

(2) Probability of False Positive (P_{fp}^{res}).

(3) Background Alarm Rate (BAR^{res}) or Probability of Background Alarm (P_{BA}^{res}).

b. Discrimination Stage ROC curves:

(1) Probability of Detection (P_d^{disc}).

(2) Probability of False Positive (P_{fp}^{disc}).

(3) Background Alarm Rate (BAR^{disc}) or Probability of Background Alarm (P_{BA}^{disc}).

c. Metrics:

(1) Efficiency (E).

(2) False Positive Rejection Rate (R_{fp}).

(3) Background Alarm Rejection Rate (R_{BA}).

d. Other:

(1) Probability of Detection by Size and Depth.

(2) Classification by type (i.e., 20-mm, 40-mm, 105-mm, etc.).

(3) Location accuracy.

(4) Equipment setup, calibration time and corresponding man-hour requirements.

(5) Survey time and corresponding man-hour requirements.

(6) Reacquisition/resurvey time and man-hour requirements (if any).

(7) Downtime due to system malfunctions and maintenance requirements.

1.3 STANDARD AND NONSTANDARD INERT ORDNANCE TARGETS

The standard and nonstandard ordnance items emplaced in the test areas are listed in Table 1. Standardized targets are members of a set of specific ordnance items that have identical properties to all other items in the set (caliber, configuration, size, weight, aspect ratio, material, filler, magnetic remanence, and nomenclature). Nonstandard targets are ordnance items having properties that differ from those in the set of standardized targets.

TABLE 1. INERT ORDNANCE TARGETS

Standard Type	Nonstandard (NS)
20-mm Projectile M55	20-mm Projectile M55
	20-mm Projectile M97
40-mm Grenades M385	40-mm Grenades M385
40-mm Projectile MKII Bodies	40-mm Projectile M813
BDU-28 Submunition	
BLU-26 Submunition	
M42 Submunition	
57-mm Projectile APC M86	
60-mm Mortar M49A3	60-mm Mortar (JPG)
	60-mm Mortar M49
2.75-inch Rocket M230	2.75-inch Rocket M230
	2.75-inch Rocket XM229
MK 118 ROCKEYE	
81-mm Mortar M374	81-mm Mortar (JPG)
	81-mm Mortar M374
105-mm HEAT Rounds M456	
105-mm Projectile M60	105-mm Projectile M60
155-mm Projectile M483A1	155-mm Projectile M483A
	500-lb Bomb
	M75 Submunition

JPG = Jefferson Proving Ground.
HEAT = high-explosive antitank.

SECTION 2. DEMONSTRATION

2.1 DEMONSTRATOR INFORMATION

2.1.1 Demonstrator Point of Contact (POC) and Address

POC: Mr. David Wright
(303) 236-1381

Address: U.S. Geological Survey (USGS)
Box 25046, Federal Center, M.S. 964
Denver, CO 80225-0046

2.1.2 System Description (provided by demonstrator)

The Tensor Magnetic Gradiometer System (TMGS) (fig. 1) has been reconfigured to improve its performance compared with the original system that was tested at YPG in 2003.

a. The system uses four three-axis fluxgate magnetometers.

b. This system has theoretical advantages in terms of rejection of distant noise sources and in target identification.

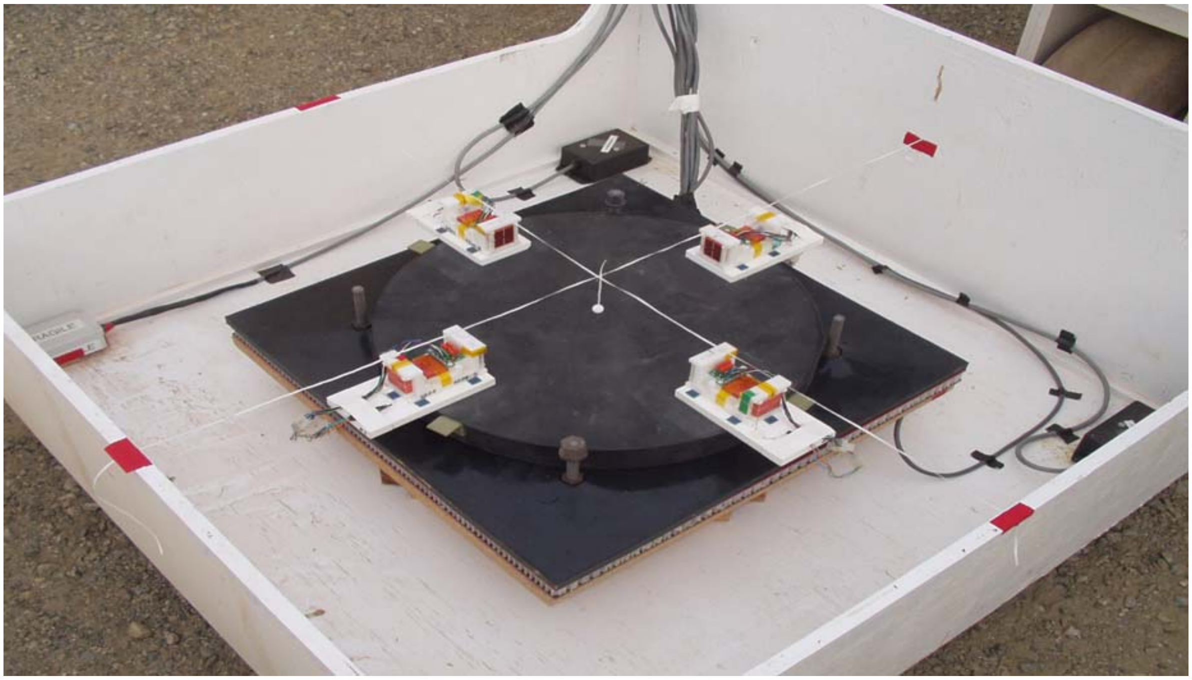

Figure 1. Demonstrator's system, TMGS MAG/towed array.

2.1.3 Data Processing Description (provided by demonstrator)

TMGS raw data files consist of an ASCII header with system settings followed by the data in binary format. GPS positions, EDA FM100B 3-axis fluxgate base station data are recorded separately on a portable PC in ASCII format and time-tagged. The Geometrics G-858G cesium vapor magnetic gradiometer logs data internally. MAGMAP2000 software transfers the data from the G-858G to a PC, where it is exported in ASCII format. The data acquisition system for the TMGS is entirely new. Details of the data format are still under development. The new system is expected to have much tighter time control, synchronized to GPS time.

2.1.4 Data Submission Format

Data were submitted for scoring in accordance with data submission protocols outlined in the Standardized UXO Technology Demonstration Site Handbook. These submitted data are not included in this report in order to protect ground truth information.

2.1.5 Demonstrator Quality Assurance (QA) and Quality Control (QC) (provided by demonstrator)

a. QC: The TMGS has real-time data displays that instantly show the operator if the transmitting/receiving functions of the system fail. In addition, plans are to find a location with no known targets and repetitively reoccupy that location and record data, including GPS data, to assess and document any drifts that may occur in the instrumentation. Standard operating procedure with all these systems is to occupy a designated clean location at least twice each day: prior to and at the completion of regular data acquisition. This usually takes place in the morning and afternoon, but in case of an extended pause in the middle of the day, an additional reference data set may be acquired. This will also test the accuracy and repeatability of the navigation data. As with all analog and time-base systems, drift will occur mainly due to component tolerances and temperature dependencies. This inherent system drift limits the absolute accuracy of the measurements that can be attained. The reference data are used primarily as a metric for overall accuracy. Abnormal drift, as would be caused by battery depletion or component degradation, would trigger a system check and data review. The hardware problem would be corrected and field data acquisition would resume. Any previous data deemed degraded would be reacquired. Plans are to preprocess data overnight or concurrent with data acquisition to visually ensure that there are no serious “glitches” or “tears” in the data. Any corrupted lines will be repeated. For the TMGS, magnetometer base stations will be set up in a magnetically clean area. These instruments monitor fluctuations in the Earth’s magnetic field, and the data will provide a check on possible sensor baseline shifts in the TMGS. Magnetic storm activity will be monitored on NOAA’s space weather web site www.sec.noaa.gov/SWN. The TMGS tetrahedron will be spin calibrated to measure deviations from orthogonality of sensor axial components. If possible, the GPS will be referenced to a local geodetic marker.

b. Line spacing for the TMGS is planned at 1 meter, based on results from 2003 and 2005 and calculations made for the new configuration. Data density along track is about 0.02 cm given a 1000 samples/second data rate.

c. The GPS positions will be plotted to assess percentage overlap and coverage. The QA and QC data can be made available along with the field data.

2.1.6 **Additional Records**

The following record(s) by this vendor can be accessed via the Internet as MicroSoft Word documents at www.uxotestsites.org.

2.2 YPG SITE INFORMATION

2.2.1 Location

YPG is located adjacent to the Colorado River in the Sonoran Desert. The UXO Standardized Test Site is located south of Pole Line Road and east of the Countermine Testing and Training Range. The open field range, calibration lanes, blind grid, mogul area, and desert extreme area comprise the 350 by 500-meter general test site area. The open field site is the largest of the test sites and measures approximately 200 by 350 meters. To the east of the open field range are the calibration and blind test grids that measure 30 by 40 meters and 40 by 40 meters, respectively. South of the open field is the 135- by 80-meter mogul area consisting of a sequence of man-made depressions. The desert extreme area is located southeast of the open field site and has dimensions of 50 by 100 meters. The desert extreme area, covered with desert-type vegetation, is used to test the performance of different sensor platforms in a more severe desert conditions/environment.

2.2.2 Soil Type

Soil samples were collected at the YPG UXO Standardized Test Site by ERDC to characterize the shallow subsurface (< 3 m). Both surface grab samples and continuous soil borings were acquired. The soils were subjected to several laboratory analyses, including sieve/hydrometer, water content, magnetic susceptibility, dielectric permittivity, X-ray diffraction, and visual description.

There are two soil complexes present within the site, Riverbend-Carrizo and Cristobal-Gunsight. The Riverbend-Carrizo complex is comprised of mixed stream alluvium, whereas the Cristobal-Gunsight complex is derived from fan alluvium. The Cristobal-Gunsight complex covers the majority of the site. Most of the soil samples were classified as either a sandy loam or loamy sand, with most samples containing gravel-size particles. All samples had a measured water content less than 7 percent, except for two that contained 11-percent moisture. The majority of soil samples had water content between 1 to 2 percent. Samples containing more than 3 percent were generally deeper than 1 meter.

An X-ray diffraction analysis on four soil samples indicated a basic mineralogy of quartz, calcite, mica, feldspar, magnetite, and some clay. The presence of magnetite imparted a moderate magnetic susceptibility, with volume susceptibilities generally greater than 100 by 10-5 SI.

For more details concerning the soil properties at the YPG test site, go to www.uxotestsites.org on the web to view the entire soils description report.

2.2.3 Test Areas

A description of the test site areas at YPG is included in Table 2.

TABLE 2. TEST SITE AREAS

Area	Description
Calibration lanes	Contains the 15 standard ordnance items buried in six positions at various angles and depths to allow demonstrator equipment calibration.
Blind grid	Contains 400 grid cells in a 0.16-hectare (0.39-acre) site. The center of each grid cell contains ordnance, clutter, or nothing.

SECTION 3. FIELD DATA

3.1 DATE OF FIELD ACTIVITIES (8 through 12 and 15 through 17 May 2006)

3.2 AREAS TESTED/NUMBER OF HOURS

Areas tested and total numbers of hours operated at each site are summarized in Table 3.

TABLE 3. AREAS TESTED AND NUMBER OF HOURS

Area	Number of Hours
Calibration lanes	32.12
Blind grid	22.35

3.3 TEST CONDITIONS

3.3.1 Weather Conditions

A YPG weather station located approximately one mile west of the test site was used to record average temperature and precipitation on a half hour basis for each day of operation. The temperatures listed in Table 4 represent the average temperature during field operations from 0700 to 1700 hours while precipitation data represents a daily total amount of rainfall. Hourly weather logs used to generate this summary are provided in Appendix B.

TABLE 4. TEMPERATURE/PRECIPITATION DATA SUMMARY

Date, 2006	Average Temperature, ^{o}F	Total Daily Precipitation, in.
8 May	89.49	0.00
9 May	88.38	0.00
10 May	90.72	0.00
11 May	93.22	0.00
12 May	95.92	0.00
15 May	94.05	0.00
16 May	92.14	0.00
17 May	99.77	0.00

3.3.2 Field Conditions

The weather was hot and the field was dry for the USGS survey.

3.3.3 Soil Moisture

Three soil probes were placed at various locations within the site to capture soil moisture data: calibration, mogul, and desert extreme areas. Measurements were collected in percent moisture and were taken twice daily (morning and afternoon) from five different soil depths (1 to 6 in., 6 to 12 in., 12 to 24 in., 24 to 36 in., and 36 to 48 in.) from each probe. Soil moisture logs are included in Appendix C.

3.4 FIELD ACTIVITIES

3.4.1 Setup/Mobilization

These activities included initial mobilization and daily equipment preparation and break down. A three-person crew took 20 hours and 11 minutes to perform the initial setup and mobilization. There was 2 hours and 16 minutes of daily equipment preparation and end of the day equipment break down lasted 50 minutes.

3.4.2 Calibration

USGS spent a total of 32 hours and 7 minutes in the calibration lanes, of which 7 hours and 2 minutes was spent collecting data.

3.4.3 Downtime Occasions

Occasions of downtime are grouped into five categories: equipment/data checks or equipment maintenance, equipment failure and repair, weather, demonstration site issues, or breaks/lunch. All downtime is included for the purposes of calculating labor costs (section 5) except for downtime due to demonstration site issues. Demonstration site issues, while noted in the daily log, are considered non-chargeable downtime for the purposes of calculating labor costs and are not discussed. Breaks and lunches are discussed in this section and billed to the total site survey area.

3.4.3.1 Equipment/data checks, maintenance. Equipment data checks and maintenance activities accounted for 4 hours and 9 minutes of site usage time. These activities included changing out batteries and routine data checks to ensure the data was being properly recorded/collected. USGS spent an additional 19 minutes for breaks and lunches.

3.4.3.2 Equipment failure or repair. No time was needed to resolve equipment failures that occurred while surveying the blind grid.

3.4.3.3 Weather. No weather delays occurred during the survey.

3.4.4 Data Collection

USGS spent a total time of 22 hours and 21 minutes in the blind grid area, 14 hours and 47 minutes of which was spent collecting data.

3.4.5 <u>Demobilization</u>

The USGS survey crew went on to conduct a full demonstration of the site. Therefore, demobilization did not occur until 17 May 2006. On that day, it took the crew 1 hour and 5 minutes to break down and pack up their equipment.

3.5 PROCESSING TIME

USGS submitted the raw data from the demonstration activities on the last day of the demonstration, as required. The scoring submittal data were provided 10 July 2006.

3.6 DEMONSTRATOR'S FIELD PERSONNEL

Mr. David L. Wright, USGS Principal Investigator
Mr. Theodore H. Asch, Research Geophysicist
Mr. Philip J. Brown, Geophysicist
Mr. Craig W. Moulton, Electronics Engineer

3.7 DEMONSTRATOR'S FIELD SURVEYING METHOD

The USGS surveyed the blind grid in a north to south direction in a linear manner.

3.8 SUMMARY OF DAILY LOGS

Daily logs capture all field activities during this demonstration and are located in Appendix D. Activities pertinent to this specific demonstration are indicated in highlighted text.

SECTION 4. TECHNICAL PERFORMANCE RESULTS

4.1 ROC CURVES USING ALL ORDNANCE CATEGORIES

Figure 2 shows the probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive. Figure 3 shows both probabilities plotted against their respective probability of background alarm. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

The overall ground truth is composed of ferrous and non-ferrous anomalies. Due to limitations of the magnetometer, the nonferrous items cannot be detected. Therefore, the ROC curves presented in this section are based on the subset of the ground truth that is solely made up of ferrous anomalies.

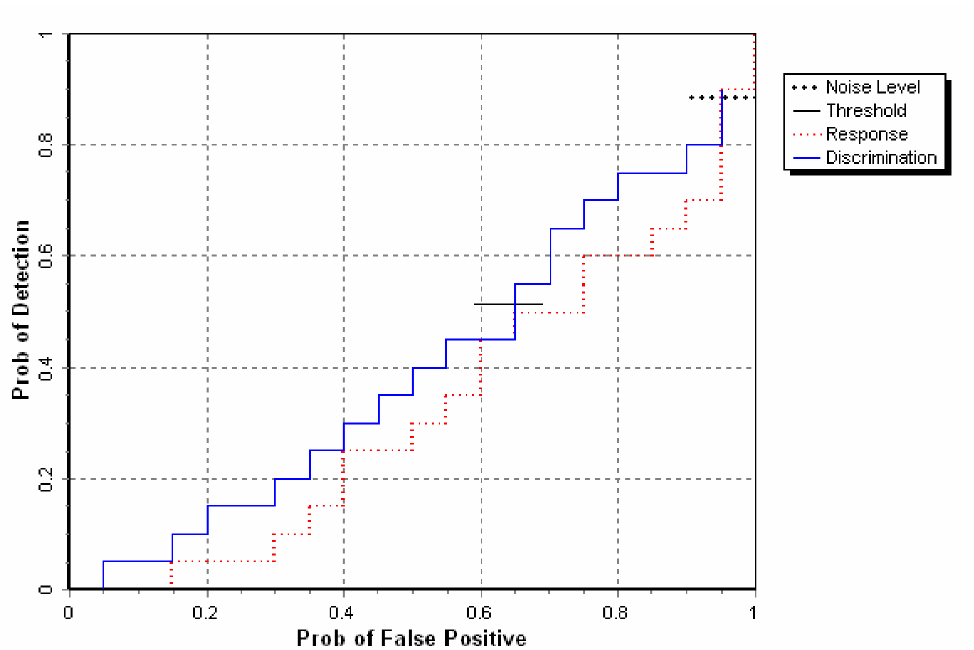

Figure 2. TMGS MAG/towed array blind grid probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined.

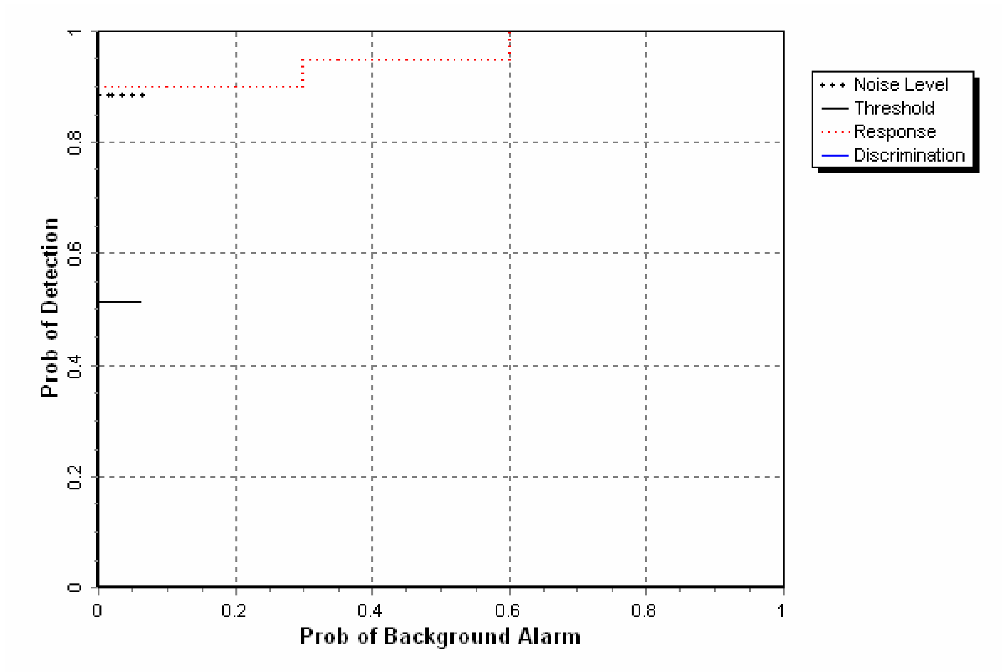

Figure 3. TMGS MAG/towed array blind grid probability of detection for response and discrimination stages versus their respective probability of background alarm over all ordnance categories combined.

4.2 ROC CURVES USING ORDNANCE LARGER THAN 20 MM

Figure 4 shows the probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive when only targets larger than 20 mm are scored. Figure 5 shows both probabilities plotted against their respective probability of background alarm. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

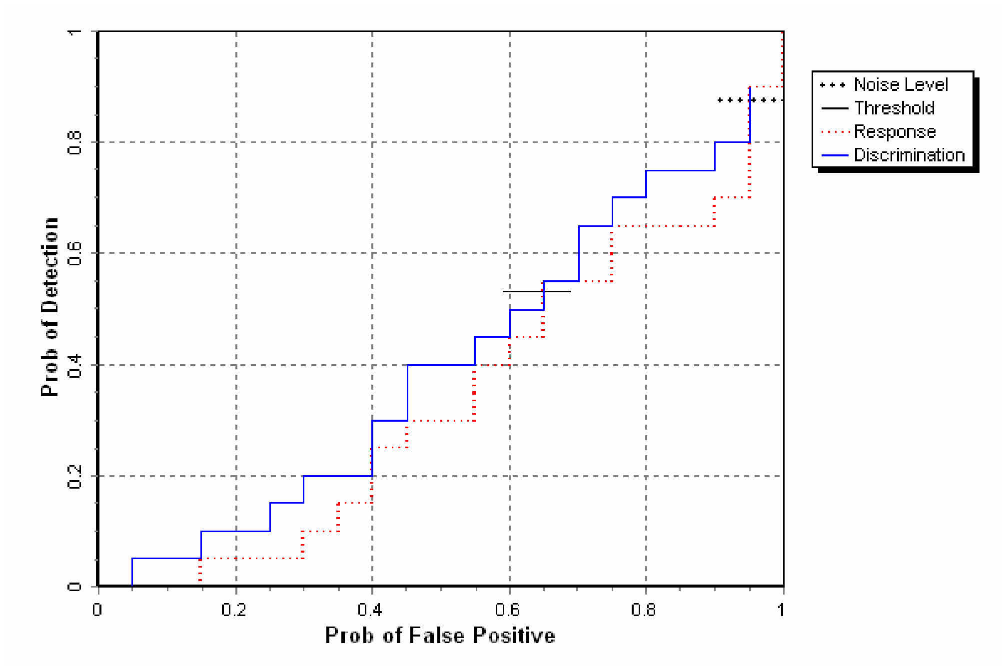

Figure 4. TMGS MAG/towed array blind grid probability of detection for response and discrimination stages versus their respective probability of false positive for all ordnance larger than 20 mm.

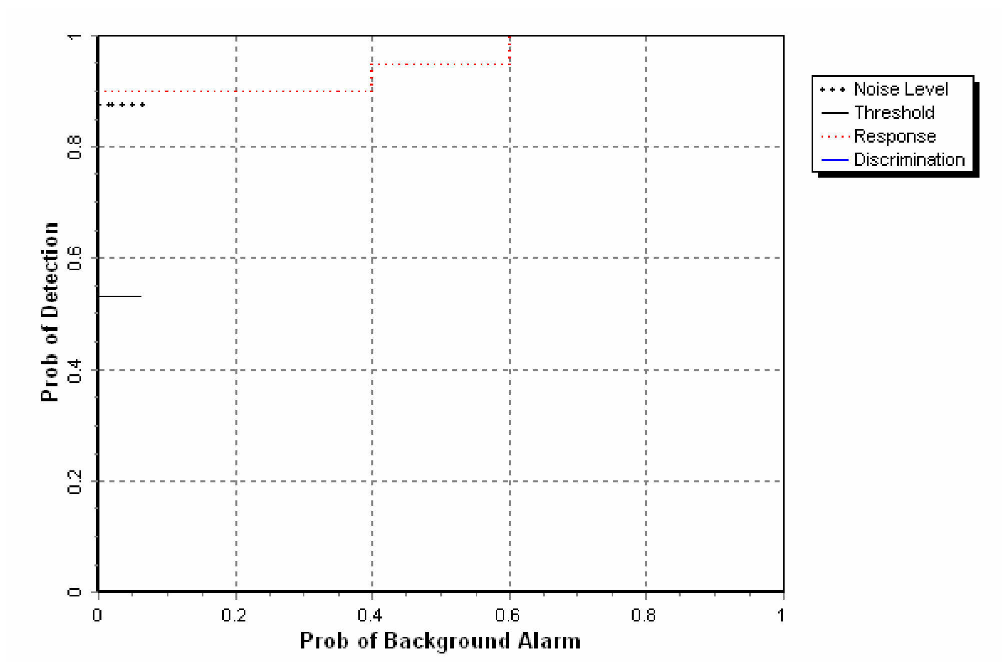

Figure 5. TMGS MAG/towed array blind grid probability of detection for response and discrimination stages versus their respective probabilities of background alarm for all ordnance larger than 20 mm.

4.3 PERFORMANCE SUMMARIES

Results for the blind grid test, broken out by size, depth and nonstandard ordnance, are presented in Tables 5a and 5b (for cost results, see section 5). Results by size and depth include both standard and nonstandard ordnance. The results by size show how well the demonstrator did at detecting/discriminating ordnance of a certain caliber range (see app A for size definitions). The results are relative to the number of ordnances emplaced. Depth is measured from the geometric center of anomalies.

The RESPONSE STAGE results are derived from the list of anomalies above the demonstrator-provided noise level. The results for the DISCRIMINATION STAGE are derived from the demonstrator's recommended threshold for optimizing UXO field cleanup by minimizing false digs and maximizing ordnance recovery. The lower 90 percent confidence limit on probability of detection and probability of false positive was calculated assuming that the number of detections and false positives are binomially distributed random variables. All results in Tables 5a and 5b have been rounded to protect the ground truth. However, lower confidence limits were calculated using actual results.

The overall ground truth is composed of ferrous and non-ferrous anomalies. Due to limitations of the magnetometer, the non-ferrous items cannot be detected. Therefore, the summary presented in Table 5a exhibits results based on the subset of the ground truth that is solely the ferrous anomalies. Table 5b exhibits results based on the full ground truth. All other tables presented in this section are based on scoring against the ferrous only ground truth. The response stage noise level and recommended discrimination stage threshold values are provided by the demonstrator.

TABLE 5a. SUMMARY OF BLIND GRID RESULTS (FERROUS ONLY)

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	≥ 1
RESPONSE STAGE									
P_d	0.90	0.85	0.95	0.90	0.80	1.00	0.95	0.90	0.45
P_d Low 90% Conf	0.82	0.77	0.81	0.81	0.63	0.85	0.90	0.76	0.17
P_d Upper 90% Conf	0.93	0.93	0.98	0.97	0.89	1.00	1.00	0.96	0.72
P_{fp}	0.95	-	-	-	-	-	0.95	0.95	N/A
P_{fp} Low 90% Conf	0.92	-	-	-	-	-	0.91	0.88	-
P_d Upper 90% Conf	0.98	-	-	-	-	-	0.98	1.00	-
P_{ba}	0.00	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.50	0.55	0.50	0.40	0.55	0.70	0.50	0.65	0.15
P_d Low 90% Conf	0.43	0.43	0.34	0.28	0.41	0.51	0.37	0.51	0.01
P_d Upper 90% Conf	0.60	0.64	0.62	0.52	0.71	0.87	0.60	0.78	0.45
P_{fp}	0.65	-	-	-	-	-	0.70	0.50	N/A
P_{fp} Low 90% Conf	0.58	-	-	-	-	-	0.62	0.37	-
P_d Upper 90% Conf	0.70	-	-	-	-	-	0.76	0.63	-
P_{ba}	0.00	-	-	-	-	-	-	-	-

Response Stage Noise Level: 50.00
Recommended Discrimination Stage Threshold: 209.00

TABLE 5b. SUMMARY OF BLIND GRID RESULTS (FULL GROUND TRUTH)

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	≥ 1
RESPONSE STAGE									
P_d	0.80	0.75	0.90	0.75	0.80	1.00	0.85	0.85	0.45
P_d Low 90% Conf	0.73	0.64	0.78	0.63	0.63	0.85	0.75	0.69	0.17
P_d Upper 90% Conf	0.86	0.82	0.96	0.83	0.89	1.00	0.91	0.91	0.72
P_{fp}	0.95	-	-	-	-	-	0.95	0.95	N/A
P_{fp} Low 90% Conf	0.92	-	-	-	-	-	0.91	0.88	-
P_d Upper 90% Conf	0.98	-	-	-	-	-	0.98	1.00	-
P_{ba}	0.00	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.45	0.45	0.50	0.35	0.55	0.70	0.40	0.60	0.15
P_d Low 90% Conf	0.39	0.36	0.35	0.24	0.41	0.51	0.33	0.47	0.01
P_d Upper 90% Conf	0.55	0.56	0.62	0.44	0.71	0.87	0.54	0.73	0.45
P_{fp}	0.65	-	-	-	-	-	0.70	0.50	N/A
P_{fp} Low 90% Conf	0.58	-	-	-	-	-	0.62	0.37	-
P_d Upper 90% Conf	0.70	-	-	-	-	-	0.76	0.63	-
P_{ba}	0.00	-	-	-	-	-	-	-	-

Response Stage Noise Level: 50.00
Recommended Discrimination Stage Threshold 209.00

4.4 EFFICIENCY, REJECTION RATES, AND TYPE CLASSIFICATION

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are reported in Table 6.

TABLE 6. EFFICIENCY AND REJECTION RATES

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	0.59	0.33	0.25
With No Loss of P_d	1.00	0.00	0.00

At the demonstrator's recommended setting, the ordnance items that were detected and correctly discriminated were further scored on whether their correct type could be identified (table 7). Correct type examples include "20-mm projectile, 105-mm HEAT projectile, and 2.75-inch rocket". A list of the standard type declaration required for each ordnance item was provided to demonstrators prior to testing. For example, the standard types for the three example items are 20mmP, 105H, and 2.75in, respectively.

TABLE 7. CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS UXO

Size	Percentage Correct
Small	7.1
Medium	23.1
Large	20.0
Overall	16.2

4.5 LOCATION ACCURACY

The mean location error and standard deviations appear in Table 8. These calculations are based on average missed depth for ordnance correctly identified in the discrimination stage. Depths are measured from the closest point of the ordnance to the surface. For the blind grid, only depth errors are calculated, since (X, Y) positions are known to be the centers of each grid square.

TABLE 8. MEAN LOCATION ERROR AND STANDARD DEVIATION (M)

	Mean	Standard Deviation
Depth	0.55	0.62

SECTION 5. ON-SITE LABOR COSTS

A standardized estimate for labor costs associated with this effort was calculated as follows: the first person at the test site was designated “supervisor”, the second person was designated “data analyst”, and the third and following personnel were considered “field support”. Standardized hourly labor rates were charged by title: supervisor at $95.00/hour, data analyst at $57.00/hour, and field support at $28.50/hour.

Government representatives monitored on-site activity. All on-site activities were grouped into one of ten categories: initial setup/mobilization, daily setup/stop, calibration, collecting data, downtime due to break/lunch, downtime due to equipment failure, downtime due to equipment/data checks or maintenance, downtime due to weather, downtime due to demonstration site issue, or demobilization. See Appendix D for the daily activity log. See section 3.4 for a summary of field activities.

The standardized cost estimate associated with the labor needed to perform the field activities is presented in Table 9. Note that calibration time includes time spent in the calibration lanes as well as field calibrations. “Site survey time” includes daily setup/stop time, collecting data, breaks/lunch, downtime due to equipment/data checks or maintenance, downtime due to failure, and downtime due to weather.

TABLE 9. ON-SITE LABOR COSTS

	No. People	Hourly Wage	Hours	Cost
Initial Setup				
Supervisor	1	$95.00	20.18	$1,917.10
Data analyst	1	57.00	20.18	1,150.26
Field support	1	28.50	20.18	575.13
Subtotal				**$3,642.49**
Calibration				
Supervisor	1	$95.00	32.12	$3,051.40
Data analyst	1	57.00	32.12	1,830.34
Field support	1	28.50	32.12	915.42
Subtotal				**$5,797.16**
Site Survey				
Supervisor	1	$95.00	22.35	$2,123.25
Data analyst	1	57.00	22.35	1,273.95
Field support	2	28.50	22.35	1,273.95
Subtotal				**$4,671.15**

See notes at end of table.

TABLE 9 (CONT'D)

	No. People	Hourly Wage	Hours	Cost
Demobilization				
Supervisor	1	$95.00	1.08	$102.60
Data analyst	1	57.00	1.08	61.56
Field support		28.50		0.00
Subtotal				$164.16
Total				**$14,274.96**

Notes: Calibration time includes time spent in the calibration lanes as well as calibration before each data run.

Site Survey time includes daily setup/stop time, collecting data, breaks/lunch, downtime due to system maintenance, failure, and weather.

SECTION 6.   COMPARISON OF RESULTS TO BLIND GRID DEMONSTRATION

No comparisons to date.

SECTION 7. APPENDIXES

APPENDIX A. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced ordnance item.

Detection: An anomaly location that is within R_{halo} of an emplaced ordnance item.

Emplaced Ordnance: An ordnance item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., non-ordnance item) buried by the government at a specified location in the test site.

R_{halo}: A pre-determined radius about the periphery of an emplaced item (clutter or ordnance) within which a location identified by the demonstrator as being of interest is considered to be a response from that item. If multiple declarations lie within R_{halo} of any item (clutter or ordnance), the declaration with the highest signal output within the R_{halo} will be utilized. For the purpose of this program, a circular halo 0.5 meters in radius will be placed around the center of the object for all clutter and ordnance items less than 0.6 meters in length. When ordnance items are longer than 0.6 meters, the halo becomes an ellipse where the minor axis remains 1 meter and the major axis is equal to the length of the ordnance plus 1 meter.

Small Ordnance: Caliber of ordnance less than or equal to 40 mm (includes 20-mm projectile, 40-mm projectile, submunitions BLU-26, BLU-63, and M42).

Medium Ordnance: Caliber of ordnance greater than 40 mm and less than or equal to 81 mm (includes 57-mm projectile, 60-mm mortar, 2.75 in. Rocket, MK118 Rockeye, 81-mm mortar).

Large Ordnance: Caliber of ordnance greater than 81 mm (includes 105-mm HEAT, 105-mm projectile, 155-mm projectile, 500-pound bomb).

Shallow: Items buried less than 0.3 meter below ground surface.

Medium: Items buried greater than or equal to 0.3 meter and less than 1 meter below ground surface.

Deep: Items buried greater than or equal to 1 meter below ground surface.

Response Stage Noise Level: The level that represents the point below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the Blind Grid test area.

Discrimination Stage Threshold: The demonstrator selected threshold level that they believe provides optimum performance of the system by retaining all detectable ordnance and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for n independent trials with the probability p of success and the probability 1-p of failure being the same for each trial. The number of successes x observed in the n trials is an estimate of p and is considered to be a binomially distributed random variable.

RESPONSE AND DISCRIMINATION STAGE DATA

The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}) and those that do not correspond to any known item, termed background alarms.

The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the RESPONSE STAGE, the demonstrator provides the scoring committee with the location and signal strength of all anomalies that the demonstrator has deemed sufficient to warrant further investigation and/or processing as potential emplaced ordnance items. This list is generated with minimal processing (e.g., this list will include all signals above the system noise threshold). As such, it represents the most inclusive list of anomalies.

The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such, and to reject clutter. For the same locations as in the RESPONSE STAGE anomaly list, the DISCRIMINATION STAGE list contains the output of the algorithms applied in the discrimination-stage processing. This list is prioritized based on the demonstrator's determination that an anomaly location is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For electronic signal processing, priority ranking is based on algorithm output. For other systems, priority ranking is based on human judgment. The demonstrator also selects the threshold that the demonstrator believes will provide "optimum" system performance, (i.e., that retains all the detected ordnance and rejects the maximum amount of clutter).

Note: The two lists provided by the demonstrator contain identical numbers of potential target locations. They differ only in the priority ranking of the declarations.

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response-stage detections)/(No. of emplaced ordnance in the test site).

Response Stage False Positive (fp^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of False Positive (P_{fp}^{res}): P_{fp}^{res} = (No. of response-stage false positives)/(No. of emplaced clutter items).

Response Stage Background Alarm (ba^{res}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): Blind Grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open Field only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{fp}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can therefore be written as $P_d^{res}(t^{res})$, $P_{fp}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to response-stage data that discriminates ordnance from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to ordnance, as well as those that the demonstrator has high confidence correspond to nonordnance or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination-stage detections)/(No. of emplaced ordnance in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm (ba^{disc}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination-stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination-stage signal strength. These quantities can therefore be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER-OPERATING CHARACERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] Figure A-1 shows how P_d versus P_{fp} and P_d versus BAR are combined into ROC curves. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

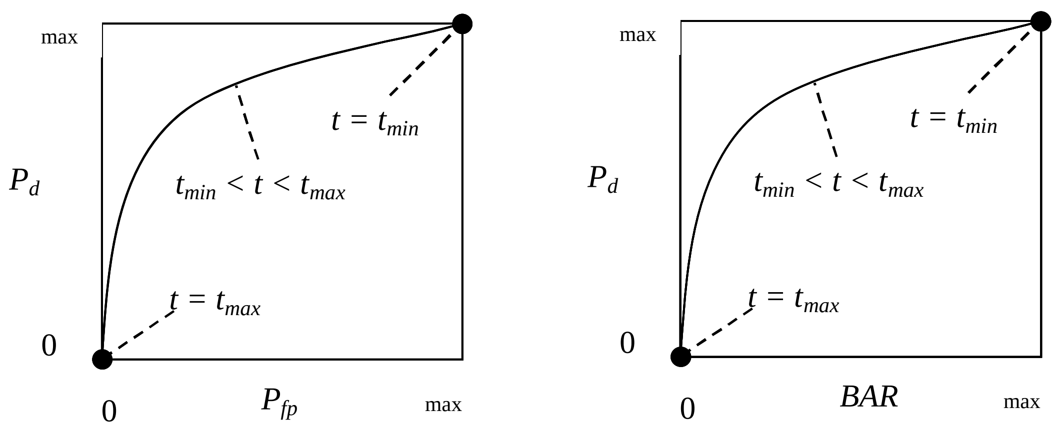

Figure A-1. ROC curves for open field testing. Each curve applies to both the response and discrimination stages.

---

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a pre-determined and fixed number of detection opportunities (some of the opportunities are located over ordnance and others are located over clutter or blank spots). In an open field scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the open field ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the Blind Grid test sites are true ROC curves.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from nonordnance items. The efficiency measures the amount of detected ordnance retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$; Measures (at a threshold of interest), the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage tmin) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the ordnance initially detected in the response stage was retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{fp}^{res}(t_{min}^{res})]$; Measures (at a threshold of interest), the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage t^{min}). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind Grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$.
Open Field: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})]$).

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON EXPLANATION:

The Chi-square test for differences in probabilities (or 2 x 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations (ref 3).

A 2 x 2 contingency table is used in the Standardized UXO Technology Demonstration Site Program to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X’s system is significantly degraded by the more challenging terrain feature introduced. The test statistic of the 2 x 2 contingency table is the

Chi-square distribution with one degree of freedom. Since an association between the more challenging terrain feature and relatively degraded performance is sought, a one-sided test is performed. A significance level of 0.05 is chosen which sets a critical decision limit of 2.71 from the Chi-square distribution with one degree of freedom. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The Chi-square test cannot be used in these instances. Instead, Fischer's test is used and the critical decision limit for one-sided tests is the chosen significance level, which in this case is 0.05. With Fischer's test, if the test statistic is less than the critical value, the proportions are considered to be significantly different.

Standardized UXO Technology Demonstration Site examples, where blind grid results are compared to those from the open field and open field results are compared to those from one of the scenarios, follow. It should be noted that a significant result does not prove a cause and effect relationship exists between the two populations of interest; however, it does serve as a tool to indicate that one data set has experienced a degradation in system performance at a large enough level than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying each of the three progressively more difficult areas using the same system (results indicate the number of ordnance detected divided by the number of ordnance emplaced):

	Blind Grid	Open Field	Moguls
P_d^{res}	100/100 = 1.0	8/10 = .80	20/33 = .61
P_d^{disc}	80/100 = 0.80	6/10 = .60	8/33 = .24

P_d^{res}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the response stage, all 100 ordnance out of 100 emplaced ordnance items were detected in the blind grid while 8 ordnance out of 10 emplaced were detected in the open field. Fischer's test must be used since a 100 percent success rate occurs in the data. Fischer's test uses the four input values to calculate a test statistic of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open field relative to results from the blind grid using the same system.

P_d^{disc}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 emplaced ordnance items were correctly discriminated as ordnance in blind grid testing while 6 ordnance out of 10 emplaced were correctly discriminated as such in open field-testing. Those four values are used to calculate a test statistic of 1.12. Since the test statistic is less than the critical value of 2.71, the two discrimination stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{res}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the response stage, 8 out of 10 and 20 out of 33 are used to calculate a test statistic of 0.56. Since the test statistic is less than the critical value of 2.71, the two response stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{disc}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the discrimination stage, 6 out of 10 and 8 out of 33 are used to calculate a test statistic of 2.98. Since the test statistic is greater than the critical value of 2.71, the smaller discrimination stage detection rate is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the ability of demonstrator X to correctly discriminate seems to have been degraded by the mogul terrain relative to results from the flat open field using the same system.

APPENDIX B. DAILY WEATHER LOGS

8 May 2006		
Time	**Temperature, ^{o}C**	**Precipitation, in.**
0700	22.3	0.00
0800	24.7	0.00
0900	27.1	0.00
1000	29.7	0.00
1100	31.7	0.00
1200	32.7	0.00
1300	33.7	0.00
1400	34.3	0.00
1500	34.5	0.00
1600	34.7	0.00
1700	34.4	0.00
9 May 2006		
Time	**Temperature, ^{o}C**	**Precipitation, in.**
0700	22.9	0.00
0800	25.0	0.00
0900	27.2	0.00
1000	29.0	0.00
1100	30.1	0.00
1200	31.1	0.00
1300	32.4	0.00
1400	33.3	0.00
1500	34.1	0.00
1600	34.4	0.00
1700	34.8	0.00
10 May 2006		
Time	**Temperature, ^{o}C**	**Precipitation, in.**
0700	21.6	0.00
0800	25.1	0.00
0900	27.7	0.00
1000	30.5	0.00
1100	31.3	0.00
1200	32.8	0.00
1300	34.0	0.00
1400	35.6	0.00
1500	37.3	0.00
1600	36.5	0.00
1700	36.6	0.00

11 May 2006		
Time	**Temperature, °C**	**Precipitation, in.**
0700	23.6	0.00
0800	26.8	0.00
0900	29.2	0.00
1000	31.5	0.00
1100	32.8	0.00
1200	34.4	0.00
1300	35.6	0.00
1400	36.5	0.00
1500	37.3	0.00
1600	37.9	0.00
1700	37.5	0.00
12 May 2006		
Time	**Temperature, °C**	**Precipitation, in.**
0700	24.0	0.00
0800	27.8	0.00
0900	31.3	0.00
1000	33.3	0.00
1100	34.6	0.00
1200	35.7	0.00
1300	36.9	0.00
1400	37.9	0.00
1500	38.8	0.00
1600	38.7	0.00
1700	38.8	0.00
15 May 2006		
Time	**Temperature, °C**	**Precipitation, in.**
0700	25.3	0.00
0800	27.0	0.00
0900	30.0	0.00
1000	32.5	0.00
1100	33.1	0.00
1200	33.7	0.00
1300	35.7	0.00
1400	37.5	0.00
1500	37.8	0.00
1600	37.8	0.00
1700	38.1	0.00

16 May 2006		
Time	**Temperature, °C**	**Precipitation, in.**
0700	23.1	0.00
0800	25.4	0.00
0900	27.9	0.00
1000	30.2	0.00
1100	32.1	0.00
1200	34.0	0.00
1300	35.6	0.00
1400	36.6	0.00
1500	37.0	0.00
1600	37.3	0.00
1700	37.3	0.00
17 May 2006		
Time	**Temperature, °C**	**Precipitation, in.**
0700	27.3	0.00
0800	30.2	0.00
0900	34.4	0.00
1000	36.1	0.00
1100	37.1	0.00
1200	37.9	0.00
1300	38.4	0.00
1400	39.7	0.00
1500	39.8	0.00
1600	39.9	0.00
1700	40.1	0.00

APPENDIX C. SOIL MOISTURE

Date: 8 May 2006			
Time: 0800 - 1330 hours			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	1.7	1.6
	6 to 12	2.3	2.3
	12 to 24	3.7	3.8
	24 to 36	3.7	3.7
	36 to 48	4.2	4.2
Mogul area	0 to 6	1.8	1.7
	6 to 12	3.7	3.7
	12 to 24	3.8	3.8
	24 to 36	4.9	4.8
	36 to 48	5.1	5.2
Desert extreme area	0 to 6	4.0	3.8
	6 to 12	3.8	3.7
	12 to 24	3.2	3.2
	24 to 36	4.1	4.0
	36 to 48	4.0	4.0
Date: 9 May 2006			
Time: 0730 - 1300 hours			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	1.8	1.5
	6 to 12	2.3	2.2
	12 to 24	3.8	3.8
	24 to 36	3.7	3.8
	36 to 48	4.2	4.2
Mogul area	0 to 6	1.8	1.9
	6 to 12	9.4	3.8
	12 to 24	3.8	3.8
	24 to 36	4.9	4.8
	36 to 48	5.2	5.8
Desert extreme area	0 to 6	1.6	1.4
	6 to 12	1.7	1.7
	12 to 24	3.4	3.3
	24 to 36	4.1	4.1
	36 to 48	4.1	4.1

Date: 10 May 2006			
Time: 0700 - 1400 hours			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	1.5	1.4
	6 to 12	2.2	2.3
	12 to 24	3.8	3.8
	24 to 36	3.7	3.7
	36 to 48	4.2	4.1
Mogul area	0 to 6	1.7	1.9
	6 to 12	3.8	3.8
	12 to 24	3.8	3.8
	24 to 36	4.9	4.9
	36 to 48	5.2	5.2
Desert extreme area	0 to 6	6.1	3.8
	6 to 12	3.3	3.8
	12 to 24	3.3	3.3
	24 to 36	4.1	4.1
	36 to 48	4.1	4.1
Date: 11 May 2006			
Time: 0730 - 1330 hours			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	1.5	1.4
	6 to 12	2.2	2.3
	12 to 24	3.8	3.7
	24 to 36	3.8	3.8
	36 to 48	4.2	4.1
Mogul area	0 to 6	1.8	1.7
	6 to 12	3.8	3.8
	12 to 24	3.8	3.8
	24 to 36	4.9	4.8
	36 to 48	5.3	5.0
Desert extreme area	0 to 6	5.9	5.3
	6 to 12	3.4	3.8
	12 to 24	3.3	3.3
	24 to 36	4.1	4.1
	36 to 48	4.1	4.1

Date: 12 May 2006			
Time: 0730 - 1300 hours			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	1.5.	1.3
	6 to 12	2.3	2.3
	12 to 24	3.7	3.7
	24 to 36	3.8	3.8
	36 to 48	4.2	4.1
Mogul area	0 to 6	1.8	1.8
	6 to 12	3.7	3.8
	12 to 24	3.8	3.8
	24 to 36	4.8	4.6
	36 to 48	5.2	5.0
Desert extreme area	0 to 6	6.0	4.9
	6 to 12	3.8	3.7
	12 to 24	3.2	3.2
	24 to 36	4.1	4.1
	36 to 48	4.0	4.1
Date: 15 May 2006			
Time: 0800 - 1330 hours			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	1.7	1.5
	6 to 12	2.0	2.1
	12 to 24	3.8	3.8
	24 to 36	3.8	3.8
	36 to 48	4.2	4.2
Mogul area	0 to 6	1.8	1.7
	6 to 12	3.8	3.8
	12 to 24	5.7	3.8
	24 to 36	4.9	4.8
	36 to 48	5.3	5.1
Desert extreme area	0 to 6	3.8	4.0
	6 to 12	3.8	3.8
	12 to 24	3.3	3.2
	24 to 36	4.1	4.0
	36 to 48	4.1	4.0

Date: 16 May 2006			
Time: 0730 - 1300 hours			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	1.7	1.6
	6 to 12	2.2	2.2
	12 to 24	3.7	3.8
	24 to 36	3.8	3.8
	36 to 48	4.2	4.2
Mogul area	0 to 6	1.8	1.8
	6 to 12	3.9	3.7
	12 to 24	3.8	3.8
	24 to 36	4.9	4.8
	36 to 48	5.1	5.0
Desert extreme area	0 to 6	3.9	3.7
	6 to 12	3.8	3.8
	12 to 24	3.3	3.2
	24 to 36	4.0	4.0
	36 to 48	4.1	4.0
Date: 17 May 2006			
Time: 0800 - 1400 hours			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	1.5	1.4
	6 to 12	2.2	2.3
	12 to 24	3.7	3.7
	24 to 36	3.8	3.8
	36 to 48	4.2	4.2
Mogul area	0 to 6	1.7	1.8
	6 to 12	3.8	3.8
	12 to 24	3.8	3.8
	24 to 36	4.8	4.8
	36 to 48	5.2	5.1
Desert extreme area	0 to 6	4.0	3.9
	6 to 12	3.6	3.7
	12 to 24	3.2	3.2
	24 to 36	4.1	4.1
	36 to 48	4.1	4.0

Date, 2006	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Track Method=Other Explain	Pattern	Field Conditions	
8 May	3	CALIBRATION LANES	0745	1127	222	INITIAL SETUP	Setting up test equipment. Unit TMGS.	NA	NA	NA	Clear	Cool
8 May	3	CALIBRATION LANES	1127	1141	14	BREAK/LUNCH	Lunch	NA	NA	NA	Sunny	Hot
8 May	3	CALIBRATION LANES	1141	1609	268	INITIAL SETUP	Continued setting up test equipment. Unit TMGS.	NA	NA	NA	Sunny	Hot
8 May	3	CALIBRATION LANES	1609	1621	12	DAILY START, STOP	Breakdown end of day	NA	NA	NA	Sunny	Hot
9 May	3	CALIBRATION LANES	0633	1121	288	DAILY START, STOP	Continued setting up test equipment. Unit TMGS.	NA	NA	NA	Clear	Cool
9 May	3	CALIBRATION LANES	1121	1155	34	BREAK/LUNCH	Lunch	NA	NA	NA	Sunny	Hot
9 May	3	CALIBRATION LANES	1155	1541	226	INITIAL SETUP	Continued setting up and calibrating test equipment. Unit TMGS.	NA	NA	NA	Sunny	Hot
9 May	3	CALIBRATION LANES	1541	1615	30	DAILY START, STOP	Breakdown end of day	NA	NA	NA	Sunny	Hot
10 May	3	CALIBRATION LANES	0637	1348	431	INITIAL SETUP	Continued setting up and calibrating test equipment. Unit TMGS.	NA	NA	NA	Sunny	Hot
10 May	3	CALIBRATION LANES	1348	1424	36	BREAK/LUNCH	Lunch	NA	NA	NA	Sunny	Hot
10 May	3	CALIBRATION LANES	1424	1528	64	INITIAL SETUP	Continued setting up and calibrating test equipment. Unit TMGS.	NA	NA	NA	Sunny	Hot
10 May	3	CALIBRATION LANES	1528	1600	32	DAILY START, STOP	Breakdown end of day	NA	NA	NA	Sunny	Hot

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date, 2006	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Track Method=Other Explain	Pattern	Field Conditions	
11 May	3	CALIBRATION LANES	0646	0824	98	DAILY START, STOP	Setting up test equipment. Unit TMGS.	NA	NA	NA	Sunny	Cool
11 May	3	CALIBRATION LANES	0824	1032	128	COLLECTING DATA	Ran Calibration grid north to south, west to east using the EM61-MK2 system	GPS	NA	Linear	Sunny	Hot
11 May	3	CALIBRATION LANES	1032	1212	100	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Downloading data from the Calibration grid and checking data	NA	NA	NA	Sunny	Hot
11 May	3	CALIBRATION LANES	1212	1249	37	BREAK/LUNCH	Lunch	NA	NA	NA	Sunny	Hot
11 May	3	CALIBRATION LANES	1249	1402	73	COLLECTING DATA	Collecting data using the TMGS	GPS	NA	Linear	Sunny	Hot
11 May	3	CALIBRATION LANES	1402	1446	44	DAILY START, STOP	Breakdown end of day	NA	NA	NA	Sunny	Hot
12 May	3	CALIBRATION LANES	0640	0925	165	DAILY START, STOP	Setting up test equipment. Unit G-858 and EDA.	NA	NA	NA	Sunny	Cool
12 May	3	CALIBRATION LANES	0925	1109	104	COLLECTING DATA	Ran Calibration grid north to south, west to east using the G-858 system	GPS	NA	Linear	Sunny	Hot
12 May	3	CALIBRATION LANES	1109	1256	107	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Downloading data from the Calibration grid and checking data	NA	NA	NA	Sunny	Hot
12 May	3	CALIBRATION LANES	1256	1326	30	BREAK/LUNCH	Lunch	NA	NA	NA	Sunny	Hot

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date, 2006	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Track Method=Other Explain	Pattern	Field Conditions	
12 May	3	CALIBRATION LANES	1326	1411	45	COLLECTING DATA	Repeating the Calibration grid in the same manner using the G-858 system, due to GPS and the system not set at the same data transfer speed	GPS	NA	Linear	Sunny	Hot
12 May	3	CALIBRATION LANES	1411	1457	46	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Downloading data from the Calibration grid and checking data	NA	NA	NA	Sunny	Hot
12 May	3	CALIBRATION LANES	1457	1546	49	COLLECTING DATA	Continued to redo the Calibration grid	GPS	NA	Linear	Sunny	Hot
12 May	3	CALIBRATION LANES	1546	1629	43	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Downloading data from the Calibration grid and checking data	NA	NA	NA	Sunny	Hot
12 May	3	CALIBRATION LANES	1629	1635	6	COLLECTING DATA	Continued to redo the Calibration grid; completed	GPS	NA	Linear	Sunny	Hot
12 May	3	CALIBRATION LANES	1635	1710	35	DAILY START, STOP	Breakdown end of day	NA	NA	NA	Sunny	Hot
15 May	4	CALIBRATION LANES	0637	1044	247	DAILY START, STOP	Setting up test equipment. Unit TMGS on the tractor.	NA	NA	NA	Sunny	Cool
15 May	4	CALIBRATION LANES	1044	1101	17	COLLECTING DATA	Running a couple of lanes on and off of the Calibration grid to test system.	GPS	NA	Linear	Sunny	Warm
15 May	4	CALIBRATION LANES	1101	1111	10	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Downloading and inspecting data	NA	NA	NA	Sunny	Warm
15 May	4	CALIBRATION LANES	1111	1154	43	BREAK/LUNCH	Lunch	NA	NA	NA	Sunny	Warm

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date, 2006	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Track Method=Other Explain	Pattern	Field Conditions	
15 May	4	CALIBRATION LANES	1154	1248	54	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Inspecting data	NA	NA	NA	Sunny	Warm
15 May	4	BLIND TEST GRID	1248	1657	249	COLLECTING DATA	Ran Blind Grid north to south, west to east using the towed TMGS system; incomplete	GPS	NA	Linear	Sunny	Warm
15 May	4	BLIND TEST GRID	1657	1713	16	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Downloading data	NA	NA	NA	Sunny	Warm
15 May	4	BLIND TEST GRID	1713	1730	17	DAILY START, STOP	Breakdown end of day	NA	NA	NA	Sunny	Warm
16 May	4	BLIND TEST GRID	0633	0753	80	DAILY START, STOP	Setting up test equipment	NA	NA	NA	Sunny	Cool
16 May	4	BLIND TEST GRID	0753	0836	43	COLLECTING DATA	Continued to run Blind Grid north to south, west to east using the towed TMGS system; completed	GPS	NA	Linear	Sunny	Warm
16 May	4	BLIND TEST GRID	0836	0859	23	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Downloading data	NA	NA	NA	Sunny	Warm
16 May	4	BLIND TEST GRID	0859	1321	262	COLLECTING DATA	Repeating the Blind Grid north to south, west to east using the towed TMGS system to verify the data taken the previous day; completed	GPS	NA	Linear	Sunny	Hot
16 May	4	BLIND TEST GRID	1321	1358	37	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Downloading data	NA	NA	NA	Sunny	Hot

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date, 2006	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Track Method=Other Explain	Pattern	Field Conditions	
16 May	4	BLIND TEST GRID	1358	1417	19	BREAK/LUNCH	Lunch	NA	NA	NA	Sunny	Hot
16 May	4	BLIND TEST GRID	1417	1422	5	COLLECTING DATA	TMGS ran surveys over selected targets outside the Calibration and Blind grids	GPS	NA	Linear	Sunny	Hot
16 May	4	BLIND TEST GRID	1422	1429	7	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Downloading data	NA	NA	NA	Sunny	Hot
16 May	4	BLIND TEST GRID	1429	1459	30	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Setting up the TMGS for the Thermal Drift test	NA	NA	NA	Sunny	Hot
16 May	4	BLIND TEST GRID	1459	1512	13	COLLECTING DATA	Collecting data on the Thermal Static test	GPS	NA	Linear	Sunny	Hot
16 May	4	BLIND TEST GRID	1512	1545	33	DAILY START, STOP	Breakdown end of day	NA	NA	NA	Sunny	Hot
17 May	4	BLIND TEST GRID	0646	0742	56	DAILY START, STOP	Setting up test equipment	NA	NA	NA	Sunny	Cool
17 May	4	BLIND TEST GRID	0742	1230	288	COLLECTING DATA	Performing the Thermal Drift test/completed	GPS	NA	Linear	Slightly Cloudy	Warm
17 May	4	BLIND TEST GRID	1230	1244	14	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Downloading data	NA	NA	NA	Sunny	Hot
17 May	4	BLIND TEST GRID	1244	1440	116	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Disassembling the towed TMGS system and reassembling the TMGS apparatus to run another spin calibration	NA	NA	NA	Sunny	Hot
17 May	2	BLIND TEST GRID	1440	1507	27	COLLECTING DATA	Performing the Spin Calibration test/completed	GPS	NA	Linear	Sunny	Hot

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date, 2006	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Track Method=Other Explain	Pattern	Field Conditions	
17 May	2	BLIND TEST GRID	1507	1513	6	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Downloading data	NA	NA	NA	Sunny	Hot
17 May	2	BLIND TEST GRID	1513	1618	65	DEMOBILIZATION	Disassembling the TMGS system	NA	NA	NA	Sunny	Hot

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

APPENDIX E. REFERENCES

1. Standardized UXO Technology Demonstration Site Handbook, DTC Project No. 8-CO-160-000-473, Report No. ATC-8349, March 2002.

2. Aberdeen Proving Ground Soil Survey Report, October 1998.

3. Data Summary, UXO Standardized Test Site: APG Soils Description, May 2002.

4. Yuma Proving Ground Soil Survey Report, May 2003.

5. Practical Nonparametric Statistics, W.J. Conover, John Wiley & Sons, 1980, pages 144 through 151.

APPENDIX F. ABBREVIATIONS

APG	=	Aberdeen Proving Ground
ASCII	=	American Standard Code for Information Interchange
ATC	=	U.S. Army Aberdeen Test Center
ERDC	=	U.S. Army Corps of Engineers Engineering Research and Development Center
ESTCP	=	Environmental Security Technology Certification Program
EQT	=	Army Environmental Quality Technology Program
GPS	=	Global Positioning System
JPG	=	Jefferson Proving Ground
NOAA	=	National Oceanic & Atmospheric Administration
P_{ba}	=	probability of background alarm rate
PC	=	personal computer
P_d	=	probability of detection
P_d^{disc}	=	probability of detection, discrimination stage
P_d^{res}	=	probability of detection, response stage
P_{fp}	=	probability of false positive
P_{fp}^{disc}	=	probability of false positive, discrimination stage
P_{fp}^{res}	=	probability of false positive, response stage
POC	=	point of contact
QA	=	quality assurance
QC	=	quality control
ROC	=	receiver-operating characteristic
SERDP	=	Strategic Environmental Research and Development Program
TMGS	=	Tensor Magnetic Gradiometer System
USAEC	=	U.S. Army Environmental Command
USGS	=	U.S. Geological Survey
UXO	=	unexploded ordnance
YPG	=	U.S. Army Yuma Proving Ground

APPENDIX G. DISTRIBUTION LIST

DTC Project No.8-CO-160-UXO-021

Addressee	No. of Copies
Commander U.S. Army Environmental Command ATTN: IMAE-ATT (Ms. Kimberly Watts) Aberdeen Proving Ground, MD 21010-5401	1
Booz Allen Hamilton ATTN: Mr. Patrick McDonnell 4692 Millennium Drive, Suite 200 Belcamp, MD 21017-1535	1
U.S. Geological Survey ATTN: Mr. David Wright Box 25046, Federal Center, M.S. 964 Denver, CO 80225-0046	1
Commander U.S. Army Aberdeen Test Center ATTN: CSTE-DTC-AT-SL-E (Mr. Michael Karwatka)	1
(Library)	1
CSTE-DTC-AT-CS-R	1
Aberdeen Proving Ground, MD 21005-5059	
Defense Technical Information Center 8725 John J. Kingman Road, Suite 0944 Fort Belvoir, VA 22060-6218	PDF

Secondary distribution is controlled by Commander, U.S. Army Environmental Command, ATTN: IMAE-ATT.